RECEIVED
CENTRAL FAX CENTER
FEB 14 2006

Certificate of Facsimile Transmission

I certify that this document, consisting of 29 pages (including this cover sheet, a response to the notice of 12/21/05, a petition for extension of time under 37 CFR 1.136(a), and a revised appeal brief), is being transmitted via facsimile to the United States Patent and Trademark Office at fax telephone number (571)-273-8300 on 2/14/06.

A E Barschall

Anne E. Barschall

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Patent Application Ser. No.: 09/822,121 Group Art Unit: 2643

Filing Date: 3/30/2001 Examiner: WING F. CHAN

Attorney Docket Number PH-US 010080 Inventor Name(s): COLMENAREZ ET AL.

Confirmation #: 8881

Title: METHOD AND APPARATUS FOR AUDIO/IMAGE SPEAKER DETECTION AND LOCATOR

Mail Stop Appeal Brief
Commissioner for Patents
P.O. Box 1450
Alexandria VA 22313-1450

RESPONSE TO FIRST NOTICE

OF NON-COMPLIANCE

Sir:

The notice of December 21, 2005 is respectfully traversed. 37 CFR 41.68 (ix) and (x) state that there is to be an appendix with any evidence or related decision. The undersigned understands this language to mean that, where there is no evidence or related decision, there should be no appendix. The requirement to add blank appendices seems strange, annoying, and an unlikely interpretation of the regulations. Withdrawal of the first notice is accordingly respectfully requested.

Nevertheless, in effort to advance prosecution, a revised appeal brief with blank appendices is included.

Also enclosed is a petition for extension of time. This petition should be considered provisional in nature. In other words, if the PTO determines that does withdraw the first notice, the extension fee should be re-credited to the deposit account.

Respectfully submitted,

CERTIFICATE OF MAILING

I hereby certify that this correspondence is being deposited this date with the United States Postal Service as first class mail in an envelope addressed to

Mail Stop Appeal Brief
Commissioner for Patents
P.O. Box 1450
Alexandria VA 22313-1450

On _____ (Mailing Date)
By _____ (Signature)

By /A Barschall/
Anne E. Barschall
Reg. No. 31,089
(914) 332-1019
fax 914-332-7719
February 14, 2006

RECEIVED
CENTRAL FAX CENTER
FEB 14 2006

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Patent Application Ser. No.: 09/822,121 Group Art Unit: 2643

Filing Date: 3/30/2001 Examiner: WING F. CHAN

Attorney Docket Number PH-US 010080 Inventor Name(s): COLMENAREZ ET AL.

Confirmation #: 8881

Title: METHOD AND APPARATUS FOR AUDIO/IMAGE SPEAKER DETECTION AND LOCATOR

Mail Stop Appeal Brief
Commissioner for Patents
P.O. Box 1450
Alexandria VA 22313-1450

APPEAL BRIEF (second revised)

Sir:

This is an appeal from the final rejection of Claims 1-25.

I. REAL PARTY IN INTEREST

The real party in interest is Koninklijke Philips Electronics, N.V., a corporation of the Netherlands.

II. RELATED APPEALS AND INTERFERENCES

Applicants are not aware of any related appeals or interferences.

III. STATUS OF CLAIMS

Claims 1-8 and 10-24 stand rejected under 35 USC 103(a) over WO 99/60788 ("Potts") in view of US 6,704,048 ("Malkin") – section 8 of the final office action.

Claims 9 and 25 stand rejected under 35 USC 103(a) per the other claims and further in view of US 5,778,082 ("Chu") – section 9 of the final office action.

There are no other claims.

All rejected claims are being appealed.

IV. STATUS OF AMENDMENTS

There was a first communication under rule 116. In a first advisory action, the Examiner said that this first communication overcame all rejections over Baker. In a second advisory action, correcting the first advisory action, the Examiner said that the first communication under rule 116 could not be entered due to an error in the claim identifier for one of the claims.

In a telephone conversation dated October 19, the Examiner indicated that if Applicants corrected the claim identifier, all rejections other than paragraphs 8 and 9 of the office action would be overcome.

A second communication under rule 116 was faxed in on October 20, 2004, correcting the problem with the claim identifier, repeating the amendments to the claims, and incorporating the prior arguments by reference. Applicants therefore believe that the only parts of the office action that remain to be overcome are paragraphs 8 and 9, though Applicants have not yet received a third advisory action stating this in writing.

Accordingly, in reliance on the Examiner's statements during the telephone interview,

Applicants will only argue against sections 8 & 9 of the office action.

V. SUMMARY OF THE CLAIMED SUBJECT MATTER

Claim 1

The claimed invention relates to a video conferencing system. The system includes a stationary image pickup device (ref #210; spec p. 4, ll. 12, 18, 20; p. 5, ll. 3-10 & 19; p. 6, l. 10). The image pickup device remains motionless during operation (spec. p. 5, ll. 4, 8, 9, 19). The system also includes an audio pickup device (ref. #'s 231, 232; spec. p. 4, l. 13-14; p. 5, l. 19) for generating audio signals (ref. # 235; spec. p. 5, l. 20) representative of sound from an audio source. The system also includes means (ref. # 270; spec. p. 5, l. 20 through p. 6, l. 9) for processing the image signals and audio signals to determine a direction of the audio source relative to a reference point. The determination of direction depends at least at times on the image signals.

Claim 5

This claim recites the video conferencing system of claim 1, further comprising an electronic pan tilt zoom system for electronically manipulating the image signals to effectively provide at least one of variable pan, tilt, and zoom functions. (spec. p. 6, ll. 10-15)

Claim 2

This claim recites the video conferencing system of claim 1, wherein said processing means (ref. #270; spec. p. 5, l. 20 through p. 6, l. 9) comprises:

an audio source localization system (ref. #240, spec. p. 5, l. 21);

a computer vision person detection system (ref. #250, spec. p. 5, l. 22); and

a multimodal speaker detection system (ref. #260, spec. p. 5, l. 22 through p. 6, l. 9).

Claim 3

This claim recites the video conferencing system of claim 2, further comprising an integrated housing (ref. #110, spec. p. 6, l. 14 et. seq.) for an integrated video conferencing system incorporating the image pickup device (ref #210; spec p. 4, ll. 12, 18, 20; p. 5, ll. 3-10 & 19; p. 6, l. 10), the audio pickup device (ref. #s 231, 232; spec. p. 4, l. 13-14; p. 5, l. 19), and the processing means (ref. #270; spec. p. 5, l. 20 through p. 6, l. 9).

Claim 4

This claim recites the video conferencing system of claim 3, wherein the integrated housing (ref. #110, spec. p. 6, l. 14 et. seq.) is sized for being portable. The portability is a functional advantage.

Claim 6

This claim recites the video conferencing system of claim 1, wherein the image pickup device (ref #210; spec p. 4, ll. 12, 18, 20; p. 5, ll. 3-10 & 19; p. 6, l. 10) is a stationary camera that remains motionless during operation of the video conferencing system.

Claim 8

This claim recites the video conferencing system of claim 2, wherein the audio source localization system (ref. #240, spec. p. 5, l. 21) detects the movement of the audio source when the audio source moves relative to the reference point, and, in response to the movement, the audio source localization system causes a change in a field of view of the image pickup device (spec. p. 5, ll. 8-9). It should be noted that this claim still depends from claim 1, which states that the image pickup device is motionless.

Claim 10

This claim has recitations analogous to those of claim 1 and further recites manipulating the image signals to produce refined image signals (spec. p. 3, ll. 4-9) depending on the determined direction; and outputting the refined image signals. Examples of refined image signals are discussed in the specification further with respect to simulating pan/tilt/zoom functions as discussed with respect to claim 5.

Claim 11

This claim depends from claim 10 and contains recitations analogous to those of claims 2 and 5.

Claim 12

This claim depends from claim 10 and recites that manipulating the image signals includes varying a field of view of the image pickup device in response to the control signals. The spec uses the pan/tilt/zoom feature as an example of varying the field of view, as explained with respect to claim 5.

Claim 14

This claim depends from claim 10, but contains recitations similar to those of claim 8. The reader is referred to the summary of claim 8, above.

Claim 15

This claim recites the method of claim 13, wherein processing the image signals includes generating control signals (265, spec. p. 6, ll. 1-2) depending on the audio based direction, and manipulating the image includes electronically panning, tilting, and/or zooming said image pickup device depending on the control signals (spec. p. 6, ll. 10-15).

Claim 16

This claim recites a video conferencing system comprising:

microphones for generating audio signals representative of sound from a speaker;

a stationary video camera, remaining motionless during operation, for

generating video signals representative of a video image;

an electronic pan tilt zoom system for manipulating video images to produce the visual effects of panning, tilting, and/or zooming;

a processor for processing the video signals and the audio signals to determine a direction of a speaker relative to a reference point and supplying control signals to the electronic pan tilt zoom system for producing images that include the speaker in the field of view of the camera, the determination of direction depending at least at times on the video signals, the control signals being generated based on the determined direction of the speaker; and

a transmitter for transmitting audio and video signals for video conferencing.

Accordingly, this claim includes, *inter alia*, the stationary/motionless limitation discussed with respect to claim 1 and the pan/tilt/zoom system limitation discussed with respect to claim 5[1].

Claim 17

This claim recites the video conferencing system of claim 1, wherein at times the determination of the direction of the audio source depends on both the image signals and the audio signals (spec. p. 5, line 18 through p. 6, line 9).

---

[1] The claim is different in scope from claim 5, however, because it does not contain the limitation from claim 1 that the determination of direction depends at least at times on the image signals.

Claim 18

This claim recites the video conferencing system of claim 1, wherein the processing includes determining the movement of the audio source depending at least at times on the image signals. (spec. p. 3, l. 7; p. 5, l. 18 through p. 6, l. 9).

Claim 19

This claim recites the video conferencing system of claim 1, wherein the processing includes tracking the position of the audio source when the audio source moves, the tracking depending at least at times on the image signals (spec. p. 3, l. 7; p. 5, l. 18 through p. 6, l. 9).

Claim 20

This claim recites the video conferencing system of claim 2, wherein the computer vision person detection system detects the movement of the audio source when the audio source moves relative to the reference point, and, in response to the movement, the computer vision person detection system causes a change in a field of view of the image pickup device. (spec. p. 3, l. 7; p. 5, l. 18 through p. 6, l. 9)

Claim 21

This claim recites the method of claim 10, wherein processing the image signals further includes:

detecting the movement of the audio source when the audio source moves; and

causing electronically, in response to the movement, an [sic] variation in a field of view of the image pickup device. (spec. p. 3, l. 7; p. 5, l. 18 through p. 6, l. 9)

Claim 22

The recitations of this claim are similar to those of claim 18, except that the claim depends from claim 10, rather than claim 1.

Claim 23

The recitations of this claim are similar to those of claim 19, except that the claim depends from claim 10, rather than claim 1.

Claim 24

This claim recites a video conferencing system, but is otherwise analogous to claim 10.

VI. GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

Section 8 of the final office action is to be reviewed, per the telephone conference of October 19, wherein the Examiner indicated that the prior sections would be overcome by the amendment of October 20.

Section 9 of the office action would appear to be moot.

VII. THE ARGUMENT

Section 8 of the Final Office Action

Section 8 of the Office Action purports to reject 23 claims over a combination of references. The references are long and complex. The Potts document contains 70 pages and 25 sheets of figures. The Malkin document contains 12 columns of text and 8 sheets of drawing.[2] Therefore, in keeping with 37 C.F.R. 1.104 (c) (ii) the Examiner is required to specify what part of the references is relied upon in rejecting each claim.

Instead, the Examiner has grouped all of the rejections together, without indicating which element of which claim is rejected over which part of the references. The comments made are not clearly applicable to all claims. Accordingly, Applicants respectfully submit that the rejections are improper.

The Potts reference: claim 1

Applicants respectfully submit that paragraph 8 of the final office action fails to make a *prima facie* case of obviousness against claim 1.

The Examiner mischaracterizes Applicants' amendment of February, 2004, by stating that page 7, second & third paragraphs, makes admissions as to the Potts reference. This is false. Those paragraphs of the amendment do not refer to the Potts reference at all. Those paragraphs of the amendment refer to the summary section of Applicants' own specification at pages 2 and 3, and therefore do not characterize any prior art.

[2] The section also cites the Baker reference: however, since the Examiner has indicated that the Baker reference is withdrawn, Applicants will only discuss the Potts & Malkin references.

The WO 99/60788 document ("Potts"), which is referred to in the fourth paragraph, is only cited by the amendment for tracking the image of a moving speaker – namely for a camera in motion - not for a device that is motionless.

Accordingly, Applicants have ***not*** made the admissions as characterized by the Examiner in the Final Office Action.

Moreover, the Examiner misconstrues Applicants' claims. The Examiner alleges that the claims recite "the camera ***not*** being motionless during operation" [emphasis added]. In fact, Claim 1 recites that the image pickup device ***is*** motionless during operation. Accordingly, the Examiner's statements about pan, tilt, zoom cameras are completely irrelevant to Applicants' claim. Since Malkin is apparently cited only for this feature, Malkin, at least as applied, would also appear to be irrelevant to the claim.

In short, the Examiner wholly fails to read the elements of claim 1 on the references; and makes false and irrelevant statements about Applicants' claims and Applicants' amendment in support of the rejection.

The Board is respectfully requested to remand this application to the Examiner for issuance of a rejection that complies with 37 CFR 1.104.

Claim 5

The office action makes no explicit rejection of this claim as distinguished from the other claims; however, it appears that the Examiner misconstrues this claim. He apparently reads it as requiring a pan, tilt and zoom camera. That is not what is recited. What is recited is a system for electronically manipulating the image signals. It is this manipulation that provides pan, tilt,

and/or zoom functions. The image pickup is still the same motionless device that was cited in claim 1.

Accordingly, as mentioned above, the references cited for pan, tilt, zoom cameras are irrelevant. The Examiner has failed to indicate any reference that teaches or suggests pan, tilt, zoom functions being generating in a motionless image pickup device. Applicants accordingly respectfully submit that the Examiner has failed to meet his burden of making a *prima facie* case against this claim.

Claim 2

The Examiner fails to indicate where the references teach or suggest these features. Applicants accordingly respectfully submit that the Examiner has failed to meet his burden of making a *prima facie* case against this claim.

Claim 3

The Examiner fails to indicate where the references teach or suggest these features. Applicants accordingly respectfully submit that the Examiner has failed to meet his burden of making a *prima facie* case against this claim.

Claim 4

The Examiner fails to indicate where the references teach or suggest this feature or the resulting functional advantage. Applicants accordingly respectfully submit that the Examiner has failed to meet his burden of making a *prima facie* case against this claim.

Claim 6

The Examiner has failed to indicate any reference that teaches or suggests this feature. Applicants accordingly respectfully submit that the Examiner has failed to meet his burden of making a *prima facie* case against this claim.

Claim 8

The Examiner fails to indicate where the references teach or suggest the recited features or the resulting functional advantage. Applicants accordingly respectfully submit that the Examiner has failed to meet his burden of making a *prima facie* case against this claim.

Claims 10 & 24

The refined image signals are a functional advantage. Accordingly, the claim distinguishes even more clearly over the references than claim 1 does.

The Examiner has completely failed to indicate where these recitations or the accompanying functional advantage are taught or suggested in the references. Applicants accordingly respectfully submit that the Examiner has failed to meet his burden of making a *prima facie* case against this claim.

Claim 24 is analogous.

Claim 11

This claim depends from claim 10 and contains recitations analogous to those of claims 2 and 5. Accordingly the arguments applicable to those claims apply to this claim.

Claim 12

The Examiner fails to indicate where the references teach or suggest this feature or the functional advantage of varying a field of view. Applicants accordingly respectfully submit that the Examiner has failed to meet his burden of making a *prima facie* case against this claim.

Claim 14

This claim depends from claim 10, but contains recitations similar to those of claim 8. Arguments applicable to claim 8 are therefore applicable to claim 14.

Claim 15

The Examiner fails to indicate where the references teach or suggest this feature. Applicants accordingly respectfully submit that the Examiner has failed to meet his burden of making a *prima facie* case against this claim.

Claim 16

This claim includes, *inter alia*, the stationary/motionless limitation discussed with respect to claim 1 and the pan/tilt/zoom system limitation discussed with respect to claim 5. The rejection is accordingly deficient with respect to these limitations, as discussed before.

Claim 17

The Examiner fails to indicate where the references teach or suggest this feature. Applicants accordingly respectfully submit that the Examiner has failed to meet his burden of making a *prima facie* case against this claim.

Claim 18

The Examiner fails to indicate where the references teach or suggests this feature. Applicants accordingly respectfully submit that the Examiner has failed to meet his burden of making a *prima facie* case against this claim.

Claim 19

The Examiner fails to indicate where the references teach or suggests this feature. Applicants accordingly respectfully submit that the Examiner has failed to meet his burden of making a *prima facie* case against this claim.

Claim 20

The Examiner fails to indicate where the references teach or suggests these features. Applicants accordingly respectfully submit that the Examiner has failed to meet his burden of making a *prima facie* case against this claim. The functional advantage of changing field of view has also not been read on the references in paragraph 8 or 9.

Claim 21

The Examiner fails to indicate where the references teach or suggests these features. Applicants accordingly respectfully submit that the Examiner has failed to meet his burden of making a *prima facie* case against this claim.

Claim 22

The Examiner fails to indicate where the references teach or suggests this feature. Applicants accordingly respectfully submit that the Examiner has failed to meet his burden of making a *prima facie* case against this claim.

Claim 23

The Examiner fails to indicate where the references teach or suggests this feature. Applicants accordingly respectfully submit that the Examiner has failed to meet his burden of making a *prima facie* case against this claim.

VIII. CONCLUSION

Applicants respectfully submit that they have answered each issue raised by the Examiner and that the application is accordingly in condition for allowance. Such allowance is therefore respectfully requested.

Respectfully submitted,

By /Anne Barschall/

Anne E. Barschall
Reg. No. 31,089
(914) 332-1019
fax 914-332-7719
February 14, 2006

CERTIFICATE OF MAILING

I hereby certify that this correspondence is being deposited this date with the United States Postal Service as first class mail in an envelope addressed to

Mail Stop Appeal Brief
Commissioner for Patents
P.O. Box 1450
Alexandria VA 22313-1450

On _____ _____ _____(Mailing Date)
By _____ _____ _____(Signature)

CLAIM APPENDIX

1. (previously presented) A video conferencing system comprising:

a stationary image pickup device, remaining motionless during operation, for generating image signals representative of an image;

an audio pickup device for generating audio signals representative of sound from an audio source; and

means for processing said image signals and said audio signals to determine a direction of the audio source relative to a reference point, the determination of direction depending at least at times on the image signals.

2. (previously presented) The video conferencing system of claim 1 wherein said processing means comprises:

an audio source localization system;

a computer vision person detection system; and

a multimodal speaker detection system.

3. (previously presented) The video conferencing system of claim 2, further comprising an integrated housing for an integrated video conferencing system incorporating the image pickup device, the audio pickup device, and the processing means.

4. (original) The video conferencing system of claim 3, wherein the integrated housing is sized for being portable.

5. (previously presented) The video conferencing system of claim 1, further comprising an electronic pan tilt zoom system for electronically manipulating the image signals to effectively provide at least one of variable pan, tilt, and zoom functions.

6. (previously presented) The video conferencing system of claim 1, wherein the image pickup device is a stationary camera that remains motionless during operation of the video conferencing system.

7. (previously presented) The video conferencing system of claim 1, wherein the processing means provides control signals to an electronic pan tilt zoom system.

8. (previously presented) The video conferencing system of claim 2, wherein the audio source localization system detects the movement of the audio source when the audio source moves relative to the reference point, and, in response to the movement, the audio source localization system causes a change in a field of view of the image pickup device.

9. (previously presented) The video conferencing system of claim 1, wherein the audio pickup device is comprised of an array of two microphones.

10. (previously presented) A method comprising the steps of:

generating, at a stationary image pickup device, remaining motionless during operation, image signals representative of an image;

generating, at an audio pickup device, audio signals representative of sound from an audio source;

processing the image signals and the audio signals to determine a direction of the audio source relative to a reference point, the determination of direction depending at least at times on the image signals;

manipulating the image signals to produce refined image signals depending on the determined direction; and

outputting said refined image signals.

11. (previously presented) The method of claim 10 further comprising the steps of:

applying said audio signals to an audio source localization system;

applying said image signals to a computer vision person detection system;

processing said audio signals and said image signals with a multimodal speaker detection system to determine the direction of the audio source;

generating control signals based on the determined direction of the audio source;

applying the control signals to an electronic pan tilt zoom system to mimic the effect of at least one function of a movable camera, said function selected from the group consisting

panning, tilting, and zooming said movable camera; and

providing an output from said electronic pan tilt zoom system.

12. (previously presented) The method of claim 10, wherein manipulating the image signals includes varying a field of view of the image pickup device in response to the control signals.

13. (original) The method of claim 10, wherein processing the audio signals includes determining an audio based direction of the audio source based on the audio signals.

14. (previously presented) The method of claim 10, wherein processing the audio signals includes detecting the movement of the audio source when the audio source moves; and

manipulating the image signals includes causing electronically, in response to the movement, a variation in a field of view of the image pickup device.

15. (previously presented) The method of claim 13, wherein processing the image signals includes generating control signals depending on the audio based direction, and manipulating the image includes electronically panning, tilting, and/or zooming said image pickup device depending on the control signals.

16. (previously presented) A video conferencing system comprising:

microphones for generating audio signals representative of sound from a speaker;

a stationary video camera, remaining motionless during operation, for generating video signals representative of a video image;

an electronic pan tilt zoom system for manipulating video images to produce the visual effects of panning, tilting, and/or zooming;

a processor for processing the video signals and the audio signals to determine a direction of a speaker relative to a reference point and supplying control signals to the electronic pan tilt zoom system for producing images that include the speaker in the field of view of the camera, the determination of direction depending at least at times on the video signals, the control signals being generated based on the determined direction of the speaker; and

a transmitter for transmitting audio and video signals for video conferencing.

17. (previously presented) The video conferencing system of claim 1, wherein at times the determination of the direction of the audio source depends on both the image signals and the audio signals.

18. (previously presented) The video conferencing system of claim 1, wherein the processing includes determining the movement of the audio source depending at least at times on the image signals.

19. (previously presented) The video conferencing system of claim 1, wherein the processing includes tracking the position of the audio source when the audio source moves, the tracking

depending at least at times on the image signals.

20. (previously presented) The video conferencing system of claim 2, wherein the computer vision person detection system detects the movement of the audio source when the audio source moves relative to the reference point, and, in response to the movement, the computer vision person detection system causes a change in a field of view of the image pickup device.

21. (previously presented) The method of claim 10, wherein processing the image signals further includes:

detecting the movement of the audio source when the audio source moves; and

causing electronically, in response to the movement, an variation in a field of view of the image pickup device.

22. (previously presented) The method of claim 10, wherein the processing includes determining the movement of the audio source depending at least at times on the image signals.

23. (previously presented) The method of claim 10, wherein the processing includes tracking the position of the audio source when the audio source moves, the tracking depending at least at times on the image signals.

24. (previously presented) A video conferencing system, comprising:

a stationary image pickup device, remaining motionless during operation, for generating image signals representative of an image;

an audio pickup device for generating audio signals representative of sound from an audio source;

means for processing the image signals and the audio signals to determine a direction of the audio source relative to a reference point, the determination of direction depending at least at times on the image signals;

means for manipulating the image signals to produce refined image signals depending on the determined direction; and

an output for outputting said refined image signals.

25. (previously presented) The video conferencing system of claim 9, wherein the array of microphones includes only two microphones.

EVIDENCE APPENDIX

RELATED PROCEEDINGS APPENDIX